PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Solution structure of the splicing factor motif of the human Prp18 protein

Fahu He,[1] Makoto Inoue,[1] Takanori Kigawa,[1,2] Mari Takahashi,[1] Kanako Kuwasako,[1] Kengo Tsuda,[1] Naohiro Kobayashi,[1] Takaho Terada,[1] Mikako Shirouzu,[1] Peter Güntert,[3] Shigeyuki Yokoyama,[1,4]* and Yutaka Muto[1]*

[1] RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

[2] Tokyo Institute of Technology, 4259 Nagatsuda-cho, Midori-ku, Yokohama 226-8502, Japan

[3] Institute of Biophysical Chemistry and Frankfurt Institute of Advanced Studies, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

[4] Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Key words: NMR; pre-mRNA splicing; splicing factor hPrp18; splicing factor motif.

INTRODUCTION

Pre-mRNA splicing is catalyzed by the spliceosome machinery, composed of the U1, U2, U4/U6, and U5 small nuclear ribonucleoproteins (snRNPs), through two consecutive transesterification reactions.[1] In addition, numerous non-snRNP splicing factors are necessary for the rearrangement of the U snRNPs within the spliceosome machinery to promote the splicing reactions. Especially, a specific prolyl *cis–trans* isomerase, cyclophilin H (CypH), is involved in the rearrangement of the components of the spliceosome machinery preceding the two transesterification reactions. Corresponding to the two commitments of human CypH (hCypH), two specific interacting factors, the U4/U6-specific proteins human Prp4 (hPrp4; U4/U6-60K) and human Prp18 (hPrp18; 28 kDa protein), mediate the interactions between hCypH and the U snRNPs in the spliceosomal machinery, respectively.

The hPrp4 protein consists of an N-terminal splicing factor motif (SFM) and seven C-terminal WD-40 domains [Fig. 1(A)]. The SFM domain of hPrp4 is responsible for the direct association with hCypH, and the WD-40 repeats mediate the interactions between hPrp4 and hPrp3 (U4/U6-90K)[2] to form a stable, RNA-free trimeric subcomplex.[3–5] This subcomplex functions as the specific component of the U4/U6 and U4/U6·U5 snRNPs and accelerates the rearrangement of U snRNPs for the first esterification reaction.

On the other hand, hPrp18 contains an N-terminal SFM domain and a C-terminal Prp18 domain[6] [Fig. 1(A)]. The C-terminal Prp18 domain, whose crystal structure has been determined,[7] associates with the U5 snRNP and the Slu7 protein during splicing.[8,9] The SFM domain of hPrp18 is also necessary for the interaction with hCypH.[10] In contrast to the hPrp4/hPrp3 complex, hPrp18 can recruit hCypH by itself. Consequently, hPrp4 and/or hPrp18 may serve as a bridge to mediate the interactions between hCypH and specific splicing factors for the appropriately timed rearrangement of the spliceosome machinery in the splicing reaction.

SFM domains have exclusively been found in Prp4 and Prp18 orthologs. Interestingly, the multiple sequence

Additional Supporting Information may be found in the online version of this article.

Abbreviations: HSQC, heteronuclear single quantum coherence; NMR, nuclear magnetic resoxxnance; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy.

Grant sponsors: RIKEN Structural Genomics/Proteomics Initiative (RSGI), The National Project on Protein Structural and Functional Analyses of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), Lichtenberg Program of the Volkswagen Foundation, Japan Society for the Promotion of Science (JSPS).

*Correspondence to: Yutaka Muto or Shigeyuki Yokoyama, RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. E-mail: ymuto@gsc.riken.jp or yokoyama@biochem.s.u-tokyo.ac.jp

Received 6 September 2011; Revised 25 October 2011; Accepted 9 November 2011
Published online 5 December 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24003

alignment [Fig. 1(B)] and the phylogenetic tree analyses [Supporting Information, Fig. S1(A)] of the SFM domains in the proteins from eukaryotes revealed that the SFM domains can be divided into two subgroups, represented by the SFM domains of Prp4 and Prp18.

The first structural information about the SFM domain was obtained by an X-ray crystallographic study of the complex between hCypH and the hPrp4 SFM domain.[11] It revealed that hCypH recognizes the SFM domain of hPrp4 in a highly specific manner, through shape complementarity and electrostatic interactions. The SFM domain of hPrp4 binds to a predominantly hydrophobic cleft that is clearly distinct from the active site of hCypH, which exhibits enzymatic activities with other proteins.[11] Intriguingly, the free SFM domain of hPrp4 in solution primarily adopts a random coil conformation, in contrast to the bound state in the CypH–Prp4 (SFM) complex.[11] On the other hand, no structural information about the SFM domain of hPrp18 is currently available.

Considering that hPrp18 and hPrp4 are involved in distinct stages of the spliceosomal reaction, a structural comparison of these SFM domains could provide information for the distinction of the two subgroups. Therefore, we determined the solution structure of the SFM in hPrp18, using multidimensional NMR spectroscopy. On the basis of our results, we will discuss the similarities and differences of the SFM domains of hPrp4 and hPrp18.

MATERIALS AND METHODS

Protein sample preparation

We screened constructs with various lengths of residues flanking the SFM domain in hPrp18, using the cell-free protein production system,[12,13] and we found that SFM in hPrp18 is well folded by itself, in contrast to the previous report for the SFM of hPrp4. For the solution structural study of SFM of hPrpr18, the gene encoding the peptide including the SFM domain (fragment T60–G122) of hPrp18 was cloned into the expression vector pCR2.1 (Invitrogen), with N-terminal GSSGSSG and C-terminal SGPSSG tags. The $^{15}N/^{13}C$ doubly labeled protein was synthesized by the cell-free protein synthesis system.[12,13] For NMR structure determination, a 1.1 m*M* sample of the uniformly $^{15}N/^{13}C$ double-labeled protein was prepared in 20 m*M* ^{2}H–Tris–HCl buffer (pH 7.0), containing 100 m*M* NaCl, 1 m*M* dithiothreitol, 0.02% NaN_3, and 10% D_2O/90% H_2O.

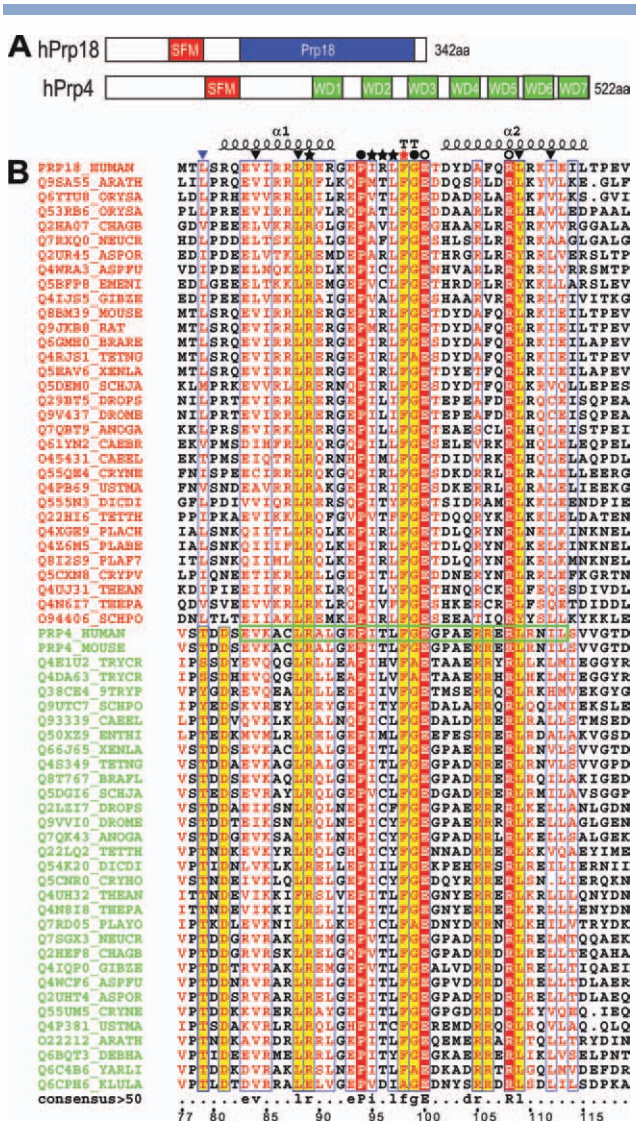

Figure 1

Domain organizations of hPrp4 and hPrp18 and structure-based multiple sequence alignment of the SFM domains. (**A**) Domain organizations of hPrp4 and hPrp18. (**B**) Structure-based multiple sequence alignment of SFM domains. The sequence alignment was produced using Clustal X, with manual correction. Protein names and corresponding species (HUMAN, *Homo sapiens*; MOUSE, *Mus musculus*; RAT, *Rattus norvegicus*; TENTG, *Tetraodon nigroviridis*; XENLA, *Xenopus laevis*; XENTR, *Xenopus tropicalis*; DROME, *Drosophila melanogaster*; DROPS, *Drosophila pseudoobscura;* ANOGA, *Anopheles gambiae*; CAEEL, *Caenorhabditis elegans*; SCHJA, *Schistosoma japonicum*; ARATH, *Arabidopsis thaliana*; ORYSA, *Oryza sativa japonica*; YEAST, *Saccharomyces cerevisiae*; KLULA, *Kluyveromyces lactis*; DEBHA, *Debaryomyces hansenii*; CANAL, *Candida albicans;* YARLI, *Yarrowia lipolytica*; SCHPO, *Schizosaccharomyces pombe*; NEUCR, *Neurospora crassa*; DICDI, *Dictyostelium discoideum AX4*; BRAFI, *Branchiostoma floridae*; ASPFU, *Aspergillus fumigatus*; CRYNE, *Cryptococcus neoformans*; USTMA, *Ustilago maydis*; PLAF7, *Plasmodium falciparum 3D7*; TENTH, *Tetrahymena thermophila*; ENTHI, *Entamoeba histolytica*; THEPA, *Theileria parva*; THEAN, *Theileria annulata*; CRYPV, *Cryptosporidium parvum*; ASHGO, *Ashbya gossypii*) are indicated following the protein accession code. The secondary structure elements are shown above the sequence alignment. The stars indicate the residues involved in intermolecular interactions (black, backbone interactions; red, side chain interactions). The open circles indicate the absolutely conserved residues with salt bridge interactions. The triangles indicate the residues forming the hydrophobic core. The blue triangle indicates a residue that is an additional contributor only in the Prp18 group. The numbering of the SFM residues in hPrp18 is shown at the bottom of the aligned sequences. The sequence fragment in the green frame is the polypeptide of the SFM in the hPrp4–hCypH complex structure. The sequence alignment was rendered with ESPript.[28]

NMR spectroscopy

All NMR experiments were performed at 298 K, using Bruker Avance 600 MHz and Avance 800 MHz spectrometers equipped with *xyz* gradient probe heads. Data were processed using the NMRPipe software package.[14] Linear prediction was used in the ^{13}C or ^{15}N dimension, to improve the digital resolution. The spectra were analyzed with the NMRView[15] and Kujira[16] programs. Backbone and side chain assignments were obtained by standard triple resonance experiments.[17] Sequence-specific backbone assignments were achieved using 2D ^{15}N–^{1}H HSQC, ^{13}C–^{1}H HSQC, 3D HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, and CBCA(CO)NH experiments. Side chain assignments were obtained using CCONH, HCCONH, HBHACONH, HCCH-COSY, and HCCH-TOCSY. Side chain NH_2 resonances of asparagine and glutamine were identified using the 3D ^{15}N–^{1}H nuclear Overhauser enhancement spectroscopy (NOESY)-HSQC spectrum.

Structure calculations

The distance restraints of the hPrp18 SFM domain for the structure calculations were obtained from the 3D ^{15}N-edited NOESY and ^{13}C-edited NOESY, with a mixing time of 100 ms. The nuclear Overhauser enhancement (NOE) cross peaks were calibrated, converted to upper limit distance restraints, and assigned, and the structures were calculated automatically with the CYANA program.[18,19] Typically, 100 structures were generated by utilizing the torsion angle dynamics protocol of CYANA, starting from random conformations.[20] Backbone ϕ and ψ dihedral angle restraints were generated from the secondary chemical shifts of the $^{1}H_{\alpha}$ and $^{13}C_{\alpha}$, $^{13}C_{\beta}$, $^{13}C'$, and ^{15}N nuclei using TALOS.[21] The 20 conformers with the lowest CYANA target function values were subjected to restrained energy-refinement with the program AMBER9, using the Generalized Born model.[22] The program MOLMOL[23] was used to analyze the resulting 20 energy-minimized conformers and to prepare drawings of the structures. The NMR structures were validated by the PROCHECK-NMR program.[24]

Complex model building

The complex model of the hCypH–hPrp18 (SFM) was built by replacing the SFM of hPrp4 in the hCypH–hPrp4 complex structure (PDB code: 1MZW) with the lowest energy solution structure of the SFM of hPrp18. The complex model was further refined by 3000 steps of energy minimization with the AMBER9 program, using the Generalized Born model.[22]

Data deposition

The 20 final structures with the lowest energy have been deposited in the Protein Data Bank, with the accession code 2DK4. The chemical shift assignments have been deposited in the BioMagResBank database, with the accession code 11355.

RESULTS AND DISCUSSION

NMR structure determination

The analytical size exclusion chromatography and the NMR spectrum revealed that the segment spanning T60-G122 of hPrp18 (UniProt accession number Q99633) is soluble and monomeric in solution [Supporting Information, Figs. S2 and S3]. We obtained complete ^{1}H, ^{15}N, and ^{13}C chemical shift assignments for the uniformly $^{13}C/^{15}N$ labeled SFM domain of hPrp18 by standard NMR methods, except for all resonances of S61, H^{N} and ^{15}N of S62, and NH_2 of N63. The NMRView[15] and Kujira[16] programs were used to obtain the sequence-specific backbone and side chain assignments. All X-Pro bonds (P76 and P94) adopted the *trans* conformation. The three-dimensional structure was determined using the combined automated NOE assignment and structure calculation algorithm of the CYANA program.[25] Experimental and analysis details are provided in the "Materials and Methods" section. The statistics of the bundle of 20 energy-refined solution conformers with the lowest CYANA target function values for the experimental restraints and the best coordinate precision and stereochemical quality, as determined by PROCHECK-NMR,[24] are shown in Table I. The structure of the hPrp18 SFM domain is well defined, based on the values of the target function, the root mean square deviation (RMSD), and the Ramachandran plot.

The solution structure of the SFM domain of hPrp18

Except for the N-terminal flexible fragment (T60-P76), the solution structure of the SFM domain of hPrp18, spanning L79-T116, contains two α helices (α1: R81–R91 and α2: D102–L115) at its N- and C-termini, respectively. The head of α1 and the tail of α2 approach each other, and the two α helices are packed against each other at an angle of about 150°. The extended, fixed loop is flanked by these helices [Fig. 2(A,B)]. A DALI search for the SFM domain of hPrp18 against known protein structures yielded only the SFM domain of hPrp4 in the complex with hCypH (PDB code: 1MZW), with a *Z*-score of 5.0. Actually, the SFM domains were identified by the well-conserved amino acid residues in the hPrp4 and hPrp18 orthologs,[10] and these SFM domains belong to Pfam PRP4 (PF08799). As described below, however, the present structural study revealed some structural

Table I
Statistics of the 20 Final Solution Structures of the hPrp18 SFM Domain

Completeness of resonance assignments	
Backbone (%)	96.8
Side chain (%)	99.2
Distance restraints	
Total NOE	1021
Intraresidue	320
Sequential ($\|i-j\|=1$)	278
Medium-range ($1<\|i-j\|<5$)	243
Long-range ($\|i-j\|\geq 5$)	180
Total dihedral angle restraints (TALOS)	
ϕ/ψ	28/27
CYANA target function (Å^2)	0.019
Structure statistics	
NOE restraint violations	
Number $>$ 0.10 Å	0
Maximum (Å)	0.12
Dihedral angle restraint violations	
Number $>$ 2.5°	0
Maximum (°)	0.10
Energies (kcal/mol)	
Mean restraint violation energy	2.91
Mean AMBER energy	−3102.07
Ramachandran plot statistics (%)[a]	
Residues in most favored regions	94.7
Residues in additional allowed regions	5.3
Residues in generously allowed regions	0
Residues in disallowed regions	0
RMSD from the average structure (Å)[a]	
Backbone atoms	0.39
Heavy atoms	1.02

[a]Ramachandran plot statistics and RMSD values are for residues L79–L115 of the SFM domain of hPrp18.

differences between the SFM domains of hPprp4 and hPrp18. A DALI search also revealed that the SAP domains[26] resemble SFM, but with larger RMSDs (>1.4 Å) and lower *Z* scores [the SAP domains fit the SFM domain of hPrp8 well with two α-helices, except for the crucial loop for the interaction; Supporting Information, Fig. S4].

In the structure of the hPrp18 SFM domain, the spatial relationship between the two α-helices is stabilized by hydrophobic interactions between residues M77 (α1), L79 (α1), V84 (α1), L88 (α1), F106 (α2), L109 (α2), R110 (α2), and I112 (α2) [Fig. 2(B)]. Among them, the positions of V84, L88, L109, and I112 are occupied by identical amino acids (V107, L111, L132, and I135, respectively) in the hPrp4 SFM domain [Fig. 1(B)], and these amino acids form the hydrophobic core in hPrp4 [Fig. 2(D)]. Furthermore, in the case of hPrpr18, the aromatic ring of F106 is surrounded by the side chains of M77 (α1), L79 (α1), and R110 (α2) to anchor the N-terminus of α1 and the C-terminus of α2 (R110 contacts F106 by a cation-π interaction). However, L79 and F106 are replaced by rather hydrophilic amino acids (T102 and R129) in hPrp4 [Fig. 1(B)]. Thus, the interactions observed in Prp18 around F106 are not present in the SFM of hPrp4 (only containing E106-L136), implying the instability of the SFM domain of Prp4 without hCypH. Besides the hydrophobic interactions, an absolutely conserved salt bridge between E100 and R108 stabilizes the α2 helix in the SFM domain in hPrp18 [Fig. 2(B)], and another between E123 and R131 reinforces the SFM of hPrpr4 [Fig. 2(D)]. However, the other salt bridge interactions for the formation of the helices in the SFM domain in hPrpr18 are missing in hPrp4. Namely, E83 and R87 could form a salt bridge in hPrp18, but this pair is replaced by E106 and S110 in hPrp4, and thus no salt bridge is expected [Fig. 2(B,D)].

As described above, in the complex structure between the SFM domain of hPrp4 and hCypH, the linker loop between helices α1 and α2 directly interacts with hCypH. The structures of the linker loop superimposed well between hPrpr4 and hPrp18 [Fig. 2(C,E)]. In the case of hPrp18, the side chains of P94 and L97 in the linker loop contact each other, and these residues are also conserved in hPrp4 [Fig. 1(B)]. However, the structures of the other linker loops are sustained differently in hPrp4 and hPrp18. Namely, the side chain of I95 on the linker loop hydrophobically interacts with that of I85 (α1). In addition, R96 on the linker loop could form a salt bridge with the side chain of D102 (α2) [Fig. 2(B)]. The structure of the linker loop is fixed by these interactions in the SFM of hPrp18. The hydrophobic interaction between I85 and I95 is specific for hPrp18. In hPrp4, I85 is replaced with K108, which could protrude into the solvent. In addition, the positions corresponding to R96 and D102, in the SFM of hPrp18, are occupied by T119 and P125 in hPrp4, respectively. Thus, these interactions are not expected in hPrpr4 [Fig. 2(D)]. On the other hand, A105 (α2) in hPrpr4 is replaced by R128 in hPrp18, and this residue could form hydrogen bonds with the carboxyl oxygen of I118 (corresponding to I95 in hPrp18) in the linker loop and the side chain of E123 (corresponding to E100 in hPrp18) in the α2 helix [Fig. 2(B,D)]. It appears that the hydrogen bonding network in the SFM of hPrpr4 was replaced by the hydrophobic and salt bridge interactions observed in the SFM of hPrpr18, to stabilize the linker loop structure.

The complex model of hCypH and the hPrp18 SFM domain

The region corresponding to the SFM in hPrp4 reportedly adopts a random coil conformation when free in solution.[11] However, the conformations of the hPrp18 SFM and the bound form of the hPrp4 SFM in the complex with hCypH are very similar, and the main chains of these domains could be superimposed on each other with an RMSD of 0.58 Å [Fig. 2(C,E)]. Thus, we built a model of the hCypH–hPrp18 (SFM) complex, on the basis of the hCypH–hPrp4 SFM complex structure. In the complex model, the F98 ring could be inserted into the hydrophobic pocket of hCypH in exactly the same man-

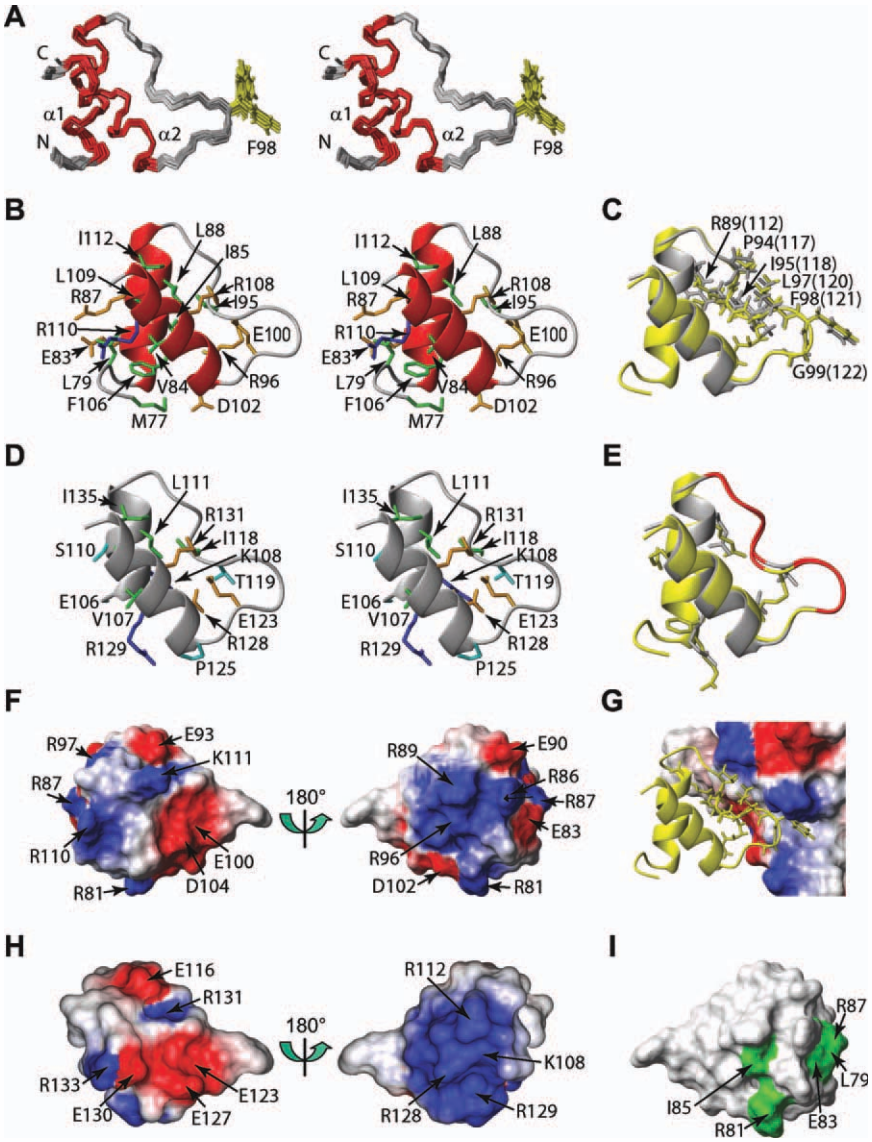

Figure 2

Tertiary structures and molecular surfaces of the SFM domains in hPrp18 and hPrp4. (**A**) Stereo view of the ensemble of 20 NMR conformers of the SFM of hPrp18. Residues from L79 to T116 are shown in a backbone wire presentation. The α1 helix (R81–R91) and the α2 helix (D102–L115) are colored red. The loops are gray. The key residue F98 for the intermolecular interactions is shown, with the side chain colored yellow. The N- and C-termini are indicated by N and C, respectively. (**B**) Ribbon representation of the lowest energy structure, showing the side chains of the conserved residues contributing to hydrophobic or salt bridge interactions. Hydrophobic residues are green; positively charged and negatively charged residues that form salt bridges are orange. The positively charged residue forming a cation-π interaction is blue. (**C**) Ribbon representation of the superimposed SFM domains of hPrp18 (gray) and hPrp4 (yellow) showing the side chains of the structurally conserved hydrophobic residues in the loop region and the residues involved in intermolecular interactions with hCypH. The numbers outside and within the parentheses of the residues are those for hPrp18 and hPrp4, respectively. (**D**) Ribbon representation of the hPrp4 SFM, with the side chains involved in salt bridges (orange) and hydrophobic interactions (green). The corresponding residues that form salt bridges and the hydrophobic core in Prp18, but are missing in Prp4, are colored cyan and blue, respectively. (**E**) Ribbon representation of the superimposed SFM domains of hPrp18 (gray) and hPrp4 (yellow), showing the same residues in the loop region (red). (**F**) Electrostatic potential surface of the SFM domain in hPrp18. (**G**) Complex model of the SFM from hPrp18 with hCypH. (**H**) Electrostatic potential surface of the SFM domain of hPrp4 (PDB code: 1MZW). The molecular surface plots are shown in the same orientation as the ribbon structures. The views on the right in H and G are rotated by 180° around the *y* axis from those on the left. (**I**) Surface representation of the conservation differences between the two SFM domain subgroups marked in Supplementary Figure 1B. The residues with conservation differences are colored green. The view is rotated by 135° around the *y* axis from that of (A).

ner as that of the SFM in hPrp4, despite the diverse side chain conformations in the 20 solution conformers of hPrp18 [Fig. 2(A,G)]. In addition, several hydrophobic amino acid residues in the linker loop (corresponding to P94, I95, L97, and F98 in hPrp18) that contact hCypH are conserved in hPrp4 and hPrp18. The main chain of

the linker loop region in the hPrp4 SFM could form hydrogen bonds with the seventh β-strand in hCypH. Thus, in the case of the hPrp18 SFM domain, the linker loop is assumed to pair with the same β-strand of hCypH, because the tertiary structures of the linker loops of hPrp4 and hPrp18 are quite similar [Fig. 2(C,E)].

The hPrp4 SFM possesses a characteristic electrostatic surface. Namely, the positively charged amino acid residues are gathered on the surface that faces hCypH and encounter a negatively charged patch on hCypH [Fig. 2(G,H)]. On the other hand, a negatively charged surface was found on the opposite side. The amino acid residues that constitute the negatively charged surface in the hPrp18 SFM (E93, E100, and D104) are located at the positions corresponding to E116, E123, and E127 in the hPrp4 SFM, respectively [left panels of Fig. 2(F,H)]. However, although both the SFM domains have positively charged surfaces on their interaction interfaces with hCypH, only one of the Arg residues on the α1 helix occupies the same position [R112 in hPrpr4, and R89 in hPrp18; right panels of Fig. 2(F,H)]. In addition to this conserved Arg residue, the positively charged surface of the hPrp18 SFM domain is composed of R81 (α1), R86 (α1), and R96 (linker loop). Thus, none of the amino acids on the α2 helix is involved in the formation of the positively charged surface of the hPrpr18 SFM. On the other hand, in the case of the hPrp4 SFM domain, K108 (α1), R128 (α2), and R129 (α2) participate in the formation of the positively charged surface, which is considered to be important for the interaction with hCypH. This is consistent with experimental results showing that hPrp4 and hPrp18 bind hCypH with high affinity.[10] Intriguingly, K108 and R129 in the hPrp4 SFM are replaced with I85 and F106 in the hPrp18 SFM, respectively, and these hydrophobic amino acid residues play an important role in the stabilization of the hPrp18 SFM domain [Fig. 2(B,D)].

Subgroup families of the SFM domains

As described above, the SFM domains were classified into two subgroups, represented by Prp4 and Prp18, respectively. We analyzed each subgroup with Hidden Markov Models,[29] using the aligned sequences in Figure 1(B) [Supporting Information, Fig. S1(B)]. Based on the sequence alignment, the residues V84, L88, R89, P94, I95, L97, F98, G99, E100, R108, L109, and I112 are strongly conserved in the subgroups [Fig. 1(B)]. As mentioned previously, some amino acid residues are important to form the hydrophobic core and to stabilize the compact fold, whereas others play a crucial role in the interaction with hCypH. Consequently, these amino acids are structurally and functionally conserved in the SFM subgroups.

However, there are still six residues that are differently conserved between the two subgroups [Supporting Information, Fig. S1(B)]. We mapped these six residues (five in α1 and one in α2) on the SFM surface [Fig. 2(I)]. These residues are in the proximity of the opposite surface of the linker loop, which is the binding site for hCypH. This suggests that this characteristic area could be responsible for the specific roles corresponding to each subgroup, besides the common binding activity for hCypH (in the case of the Prp18 subgroup, these residues stabilized the folding of the SFM domain, as described above). Prp4 primarily interacts with Prp3 through its C-terminal WD-repeats. In yeast Prp4, however, the N-terminal region of yPrp4 containing the SFM also reportedly interacts with yPrp3.[27] This implies that the SFM of yPrp4 could be stabilized by yPrp3. On the other hand, Prp18 could associate with CypH by itself. The structural stability of the Prp18 SFM, as described here, could contribute to the independence of Prp18 and suggests that different control mechanisms exist between Prp4 and Prp18.

In summary, we have presented the solution structure of the SFM of hPrp18, which forms a compact fold with two α helices interacting head-to-tail and a fixed linker loop. The solution structure of the hPrp18 SFM, which is stabilized by more interaction forces, including hydrophobic, salt-bridge, hydrogen bonding, cation-π, and so forth, than the hPrp4 SFM, is folded in solution and is quite similar to the bound state of the SFM domain in the hCypH–hPrp4 complex. The present hPrp18 SFM solution structure and the model of its complex with hCypH have shed light on the process of the second catalytic step of splicing. The distinct structural features of these two subgroups (Prp4 and Prp18) imply their different functional roles in the first and second steps of pre-mRNA splicing.

ACKNOWLEDGMENTS

We are grateful to Yasuko Tomo, Eiko Seki, Masaomi Ikari, and Yuki Kamewari-Hayami for sample preparation. This work was supported by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT).

REFERENCES

1. Will CL, Luhrmann R. Spliceosome structure and function. Cold Spring Harb Perspect Biol 2011;3:a003707.
2. Ayadi L, Miller M, Banroques J. Mutations within the yeast U4/U6 snRNP protein Prp4 affect a late stage of spliceosome assembly. RNA 1997;3:197–209.
3. Teigelkamp S, Achsel T, Mundt C, Gothel SF, Cronshagen U, Lane WS, Marahiel M, Luhrmann R. The 20 kD protein of human [U4/U6.U5] tri-snRNPs is a novel cyclophilin that forms a complex with the U4/U6-specific 60 kD and 90 kD proteins. RNA 1998;4:127–141.
4. Lauber J, Plessel G, Prehn S, Will CL, Fabrizio P, Groning K, Lane WS, Luhrmann R. The human U4/U6 snRNP contains 60 and 90 kD proteins that are structurally homologous to the yeast splicing factors Prp4p and Prp3p. RNA 1997;3:926–941.

5. Horowitz DS, Kobayashi R, Krainer AR. A new cyclophilin and the human homologues of yeast Prp3 and Prp4 form a complex associated with U4/U6 snRNPs. RNA 1997;3:1374–1387.
6. Horowitz DS, Krainer AR. A human protein required for the second step of pre-mRNA splicing is functionally related to a yeast splicing factor. Genes Dev 1997;11:139–151.
7. Jiang J, Horowitz DS, Xu RM. Crystal structure of the functional domain of the splicing factor Prp18. Proc Natl Acad Sci USA 2000;97:3022–3027.
8. Bacikova D, Horowitz DS. Mutational analysis identifies two separable roles of the *Saccharomyces cerevisiae* splicing factor Prp18. RNA 2002;8:1280–1293.
9. Horowitz DS, Abelson J. A U5 small nuclear ribonucleoprotein particle protein involved only in the second step of pre-mRNA splicing in *Saccharomyces cerevisiae*. Mol Cell Biol 1993;13:2959–2970.
10. Horowitz DS, Lee EJ, Mabon SA, Misteli T. A cyclophilin functions in pre-mRNA splicing. EMBO J 2002;21:470–480.
11. Reidt U, Wahl MC, Fasshauer D, Horowitz DS, Luhrmann R, Ficner R. Crystal structure of a complex between human spliceosomal cyclophilin H and a U4/U6 snRNP-60K peptide. J Mol Biol 2003;331:45–56.
12. Matsuda T, Kigawa T, Koshiba S, Inoue M, Aoki M, Yamasaki K, Seki M, Shinozaki K, Yokoyama S. Cell-free synthesis of zinc-binding proteins. J Struct Funct Genomics 2006;7:93–100.
13. Kigawa T, Yabuki T, Matsuda N, Matsuda T, Nakajima R, Tanaka A, Yokoyama S. Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. J Struct Funct Genomics 2004;5:63–68.
14. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293.
15. Johnson BA, Blevins RA. NMR view: a computer program for the visualization and analysis of NMR data. J Biomol NMR 1994;4:603–614.
16. Kobayashi N, Iwahara J, Koshiba S, Tomizawa T, Tochio N, Güntert P, Kigawa T, Yokoyama S. KUJIRA, a package of integrated modules for systematic and interactive analysis of NMR data directed to high-throughput NMR structure studies. J Biomol NMR 2007; 39:31–52.
17. Clore GM, Gronenborn AM. Multidimensional heteronuclear nuclear magnetic resonance of proteins. Methods Enzymol 1994;239:349–363.
18. Güntert P. Automated NMR structure calculation with CYANA. Methods Mol Biol 2004;278:353–378.
19. Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol 2002;319:209–227.
20. Güntert P, Mumenthaler C, Wüthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol 1997;273:283–298.
21. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 1999;13:289–302.
22. Case DA, Cheatham TE, 3rd Darden T, Gohlke H, Luo R, Merz KM, Jr. Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. J Comput Chem 2005;26:1668–1688.
23. Koradi R, Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 1996;14:51–55, 29–32.
24. Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 1996;8:477–486.
25. Güntert P. Automated structure determination from NMR spectra. Eur Biophys J 2009;38:129–143.
26. Aravind L, Koonin EV. SAP—a putative DNA-binding motif involved in chromosomal organization. Trends Biochem Sci 2000;25:112–114.
27. Ayadi L, Callebaut I, Saguez C, Villa T, Mornon JP, Banroques J. Functional and structural characterization of the prp3 binding domain of the yeast prp4 splicing factor. J Mol Biol 1998;284:673–687.
28. Gouet P, Courcelle E, Stuart DI, Metoz F. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 1999; 15:305–308.
29. Schuster-Bockler B, Schultz J, Rahmann S. HMM Logos for visualization of protein families. BMC Bioinform 2004;5:7.